

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 09-081174

(43)Date of publication of application : 28.03.1997

(51)Int.Cl.

G10L 3/00
G06F 17/28
G10L 5/04

(21)Application number : 07-235835

(71)Applicant : TOSHIBA CORP

(22)Date of filing : 13.09.1995

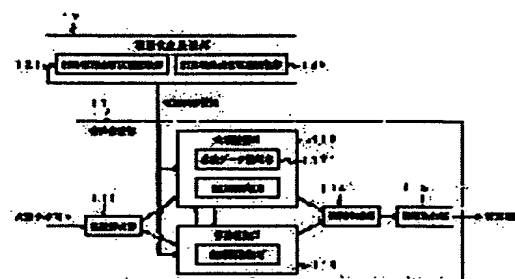
(72)Inventor : SETO SHIGENOBU
ARAI TAKAAKI

(54) VOICE SYNTHESIZING SYSTEM AND METHOD THEREFOR

(57)Abstract:

PROBLEM TO BE SOLVED: To generate such output of a system that makes a user easily grasp the user's situation and the condition of the system and to improve the operation convenience by changing an applicable phoneme/rhythm control regulation in accordance with the user's situation, user's environment and the condition of the system.

SOLUTION: The communication state in a computer in which a voice synthesizing system is operated and a communication state between the computer and the outside are monitored by a communication state monitoring part 12 and corresponding information of the communication state is outputted. In a language analysis part 111 in a voice synthesizing part 11, by performing language analysis such as morpheme analysis and syntactic structural analysis of an input text and applying regulations of various levels to the contents of the result of language analysis in a phoneme control part 112 and a rhythm control part 113, phoneme/rhythm control influencing the quality of the synthesized voice accompanied by conversion from language media to voice media is performed. At this time, the regulation applied to the phoneme/rhythm control is changed according to the information of the communication state.



(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平9-81174

(43)公開日 平成9年(1997)3月28日

(51)Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 3/00			G 1 0 L 3/00	H
G 0 6 F 17/28			5/04	N
G 1 0 L 5/04			G 0 6 F 15/38	F
				V

審査請求 未請求 請求項の数7 O L (全 16 頁)

(21)出願番号 特願平7-235835

(22)出願日 平成7年(1995)9月13日

(71)出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72)発明者 瀬戸 重宣

大阪府大阪市北区大淀中1丁目1番30号

株式会社東芝関西支社内

(72)発明者 新居 孝章

大阪府大阪市北区大淀中1丁目1番30号

株式会社東芝関西支社内

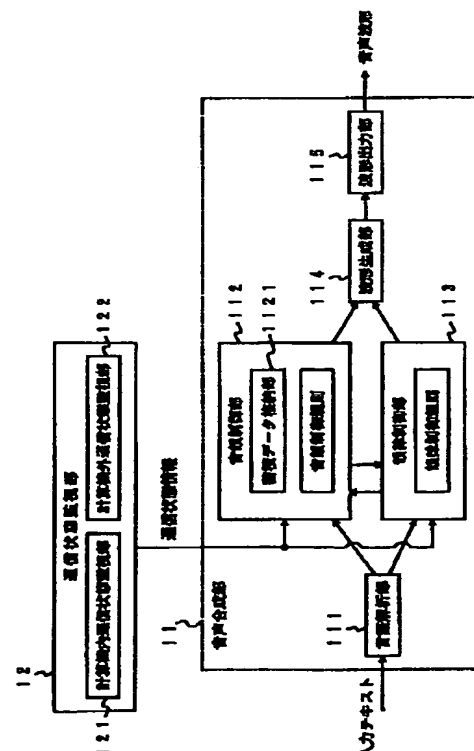
(74)代理人 弁理士 鈴江 武彦

(54)【発明の名称】 音声合成システムおよび音声合成方法

(57)【要約】

【課題】ユーザ状況やユーザ環境、システム状況に応じて、適用する音韻・韻律制御規則を変えることで、ユーザの置かれている状況やシステム状況がユーザにとって把握しやすいシステムの出力を生成し、使い勝手を向上させる。

【解決手段】音声合成システムの稼働する計算機内の通信状態、および計算機と外部の間の通信状態を通信状態監視部12にて監視し、対応する通信状態情報を出力する。音声合成部11内の言語解析部111では、入力テキストに対して形態素解析や統語構造解析などの言語解析を行い、その言語解析の結果の内容に対し、音韻制御部112および韻律制御部113にて、それぞれ様々なレベルの規則を適用することで、言語メディアから音声メディアへの変換に伴う合成音声の品質を左右する音韻・韻律的な制御を行う。この際、音韻・韻律的な制御において適用される規則が、上記の通信状態情報に応じて変更される。



【特許請求の範囲】

【請求項1】 入力テキストの言語解析を行いその解析結果に対して規則を適用して音韻・韻律的な制御を行い合成音声を生成・出力する音声合成手段と、
計算機内の通信状態、および計算機と外部の間の通信状態の少なくとも一方を監視し通信状態情報を出力する通信状態監視手段とを具備し、
前記音声合成手段は、前記音韻・韻律的な制御において適用する規則を前記通信状態監視手段から出力される前記通信状態情報に応じて変更するように構成されていることを特徴とする音声合成システム。

【請求項2】 入力テキストの言語解析を行いその解析結果に対して規則を適用して音韻・韻律的な制御を行い合成音声を生成・出力する音声合成手段と、
計算機ハードウェアの動作状態、および計算機ソフトウェアの動作状態の少なくとも一方を監視し動作状態情報を出力する動作状態監視手段とを具備し、
前記音声合成手段は、前記音韻・韻律的な制御において適用する規則を前記動作状態監視手段から出力される前記動作状態情報に応じて変更するように構成されていることを特徴とする音声合成システム。

【請求項3】 前記音声合成手段における言語解析、音韻的な制御、韻律的な制御、および音声波形生成の少なくとも1つの処理を、前記動作状態情報の示す動作状態に応じて、通信可能な有線ネットワークもしくは無線ネットワークで結ばれる複数の計算機ハードウェアに分担させることを特徴とする請求項2記載の音声合成システム。

【請求項4】 入力テキストの言語解析を行いその解析結果に対して規則を適用して音韻・韻律的な制御を行い合成音声を生成・出力する音声合成手段と、
ユーザのシステム利用状況、およびユーザのシステム利用環境の少なくとも一方を監視しユーザ状況情報を出力するユーザ状況監視手段とを具備し、
前記音声合成手段は、前記音韻・韻律的な制御において適用する規則を前記ユーザ状況監視手段から出力される前記ユーザ状況情報に応じて変更するように構成されていることを特徴とする音声合成システム。

【請求項5】 前記ユーザ状況情報をもとに人間の発声でないことを明示すべきか否かの判断結果を出力する非自然音声明示判断手段をさらに具備するとともに、
前記非自然音声明示判断手段の判断結果に応じ、前記入力テキストの表現の一部の変更により人間の音声でないことを明示するテキスト変更手段、および前記非自然音声明示判断手段の判断結果に応じ、合成音声の出力に併せて人間の発声でないことを明示する音を出力する非自然音声明示音出力手段の少なくとも一方をさらに具備することを特徴とする請求項4記載の音声合成システム。

【請求項6】 入力テキストの言語解析を行いその解析結果に対して規則を適用して音韻・韻律的な制御を行い

合成音声を生成・出力する音声合成方法であって、
計算機内の通信状態、および計算機と外部の間の通信状態の少なくとも一方、計算機ハードウェアの動作状態、および計算機ソフトウェアの動作状態の少なくとも一方、または、ユーザのシステム利用状況、およびユーザのシステム利用環境の少なくとも一方を監視し、その監視結果に応じて、前記音韻・韻律的な制御において適用する規則を変更するようにしたことを特徴とする音声合成方法。

【請求項7】 入力テキストの言語解析を行いその解析結果に対して規則を適用して音韻・韻律的な制御を行い合成音声を生成・出力する音声合成方法であって、
計算機内の通信状態、および計算機と外部の間の通信状態の少なくとも一方、計算機ハードウェアの動作状態、および計算機ソフトウェアの動作状態の少なくとも一方、または、ユーザのシステム利用状況、およびユーザのシステム利用環境の少なくとも一方を監視し、その監視結果に応じて、前記音韻・韻律的な制御において適用する規則を変更するとともに、前記入力テキストに対応する合成音に併せて別の音または合成音の少なくとも一方を出力するようにしたことを特徴とする音声合成方法。

【発明の詳細な説明】**【0001】**

【発明の属する技術分野】 本発明は、音声合成技術を利用するシステム一般に用いて好適な音声合成システムおよび音声合成方法に関する。

【0002】

【従来の技術】 近年、音声合成技術の応用が拡大され、さらに計算機の処理能力が向上するにつれ、我々の身近における音声合成の利用がますます増え、テキスト音声変換は1つの応用アプリケーションとして気軽に利用可能になってきた。岩田他：“パソコン向けソフトウェア日本語テキスト音声合成,” 日本音響学会講演論文集, 2-8-13, pp.245-246 (1993年10月) がその例である。

【0003】 これらの音声合成システムは、入力されるテキストの言語解析、音韻制御、韻律制御、波形生成のいずれの処理においても原則的に一意の処理結果を得るように作られており、入力テキストが決まれば結果として得られる合成音声は常に同じものになっていた。

【0004】 システムによっては、男声・女声、高い声・低い声など、合成音声の生成に先立ちユーザに選択させ、その選択に応じた合成音声を生成するものもあるが、選択項目が決まれば、入力テキストに対して生成される合成音声は一意に決まるという意味で、同じ枠組みであると言える。

【0005】

【発明が解決しようとする課題】 しかしながら、入力テキストに対して同じ合成音声が生産されることは、単調で飽きがくるというだけでなく、必ずしも音声メディアの特徴を有効に利用しているとは言えない。

【0006】音声メディアは、テキストメディアと異なり韻律や声色の変化が加わることにより、言語表現による直接的なメッセージだけでなく、感情や意図、話者の状況やノリといった雰囲気など、付加的な情報を副次的に伝える特徴があることはよく知られている。藤崎他：“音声の韻律的特徴による発話意図の表現，”日本音響学会講演論文集，2-8-16，pp.225-226（1993年 3月）や、上床他：“音声の感情表現の分析とモデル化，”電子情報通信学会技術研究報告，SP92-131，pp.65-72（1993年 1月）などがその例である。

【0007】上述した現状の音声合成システムやその応用アプリケーションはいずれも、文字言語メディアの形態に表現される言語情報を単に音声メディアの形態の表現に変換するメディア変換（テキスト音声変換）としての機能は持っていない、副次的な情報をも伝えるという音声の特徴を積極的に利用しているとは言い難い。

【0008】さらに、音声合成が、単独の装置としてではなく、他のシステムとの連携を行い動作する場合や、1つの応用アプリケーションとしてパーソナルコンピュータやワークステーションなどの汎用的な計算機の上で他の応用アプリケーションとともに利用される場合においても、ともに動作しているシステムやアプリケーションなどの状況や、音声合成アプリケーションが動作しているシステムがどのような状態にあるかなどの状況によらず、与えられたテキストを単に忠実に音声へ変換しているに過ぎない。

【0009】本発明は上記事情を考慮してなされたものでその目的は、システムの動作状況（システム状況）、あるいはユーザ自身の状況（ユーザ状況）やユーザのいる場所の環境（ユーザ環境）に応じて音韻・韻律制御を動的に変えたり、さらには入力テキストに対応した本来の合成音に併せて別の音や合成音を出力することで、ユーザにとってシステムの動作状況が把握しやすいようなシステムの出力や、ユーザの置かれている状況に適したシステムの出力が生成でき、使い勝手を向上させることができる音声合成システムおよび音声合成方法を提供することにある。

【0010】

【課題を解決するための手段】上述した課題を解決するため、本発明の第1の観点に係る構成は、入力テキストの言語解析を行いその解析結果に対して規則を適用して音韻・韻律的な制御を行い合成音声を生成・出力する音声合成手段に加えて、計算機内の通信状態、および計算機と外部の間の通信状態の少なくとも一方を監視し通信状態情報を出力する通信状態監視手段を備え、上記音声合成手段にあつては、上記音韻・韻律的な制御において適用する規則を通信状態監視手段から出力される通信状態情報に応じて変更するようにしたことを特徴とする。

【0011】本発明の第2の観点に係る構成は、上記第1の観点に係る構成における音声合成手段に相当する音

声合成手段に加えて、計算機ハードウェアの動作状態、および計算機ソフトウェアの動作状態の少なくとも一方を監視し動作状態情報を出力する動作状態監視手段を備え、上記音声合成手段にあつては、上記音韻・韻律的な制御において適用する規則を動作状態監視手段から出力される動作状態情報に応じて変更するようにしたことを特徴とする。ここで、上記音声合成手段における言語解析、音韻的な制御、韻律的な制御、および音声波形生成の少なくとも1つの処理を、上記動作状態情報の示す動作状態に応じて、通信可能な有線ネットワークもしくは無線ネットワークで結ばれる複数の計算機ハードウェアに分担させるようにするとよい。

【0012】本発明の第3の観点に係る構成は、上記第1の観点に係る構成における音声合成手段に相当する音声合成手段に加えて、ユーザのシステム利用状況、およびユーザのシステム利用環境の少なくとも一方を監視しユーザ状況情報を出力するユーザ状況監視手段を備え、上記音声合成手段にあつては、上記音韻・韻律的な制御において適用する規則をユーザ状況監視手段から出力されるユーザ状況情報に応じて変更するようにしたことを特徴とする。ここで、ユーザ状況情報をもとに人間の発声でないことを明示すべきか否かの判断結果を出力する非自然音声明示判断手段と、この非自然音声明示判断手段の判断結果に応じ、上記入力テキストの表現の一部の変更により人間の音声でないことを明示するテキスト変更手段、および当該判断結果に応じ、合成音声の出力に併せて人間の発声でないことを明示する音を出力する非自然音声明示音出力手段の少なくとも一方をさらに備えることも可能である。

【0013】上記第1の観点に係る構成においては、音声合成手段内で、まず入力テキストに対して形態素解析や統語構造解析などの周知の言語解析が行われて、形態素の系列に分解されるとともに「読み」を表す記号列と形態素の品詞、活用、アクセント型、形態素間の係り受け関係の強さなどの情報が併せて出力される。

【0014】また、音声合成手段内では、上記の言語解析の結果の内容に対して、それぞれ様々なレベルの規則（音韻・韻律的規則）を適用することによって、言語メディアから音声メディアへの変換に伴う合成音声の品質を左右する制御、すなわち音韻・韻律的な制御が行われ、入力テキストに対応する音声波形が生成される。

【0015】一方、通信監視手段は、音声合成システムの稼働する計算機内の通信状態、および計算機と外部の間の通信状態の少なくとも一方を監視する。それぞれの通信状態監視結果は、音声合成手段に伝えられる。

【0016】音声合成手段内では、この通信状態監視結果に応じて、上記音韻・韻律的規則の適用内容が変更される。次に、上記第2の観点に係る構成においても、音声合成手段内では、入力テキストに対する言語解析の結果の内容に対して、それぞれ様々なレベルの音韻・韻律

的規則を適用することによって、言語メディアから音声メディアへの変換に伴う合成音声の品質を左右する音韻・韻律的な制御が行われ、入力テキストに対応する音声波形が生成される。

【0017】一方、動作状態監視手段は、音声合成システムの稼働する計算機ハードウェアの動作状態、および計算機ソフトウェアの動作状態の少なくとも一方を監視する。それぞれの動作状態監視結果は、音声合成手段に伝えられる。

【0018】音声合成手段内では、この動作状態監視結果に応じて、上記音韻・韻律的規則の適用内容が変更される。また、音声合成手段における言語解析、音韻的な制御、韻律的な制御、および音声波形生成の少なくとも1つの処理が、動作状態監視結果に応じて、通信可能な有線ネットワークもしくは無線ネットワークで結ばれる複数の計算機ハードウェアに分担させられる。

【0019】次に、上記第3の観点に係る構成においても、音声合成手段内では、入力テキストに対する言語解析の結果の内容に対して、それぞれ様々なレベルの音韻・韻律的規則を適用することによって、言語メディアから音声メディアへの変換に伴う合成音声の品質を左右する音韻・韻律的な制御が行われ、入力テキストに対応する音声波形が生成される。

【0020】一方、ユーザ状況監視手段は、ユーザのシステム利用状況、およびユーザのシステム利用環境の少なくとも一方を監視する。それぞれのユーザ状況監視結果は、音声合成手段に伝えられる。

【0021】音声合成手段内では、このユーザ状況監視結果に応じて、上記音韻・韻律的規則の適用内容が変更される。また、非自然音声明示判断手段と、テキスト変更手段および非自然音声明示音出力手段の少なくとも一方とをさらに備えた構成では、入力テキストの表現の一部の変更（例えば、入力テキストに対する定型表現の追加）によって人間の音声でないことを明示するテキスト変更、あるいは合成音声の出力に併せて人間の発声でないことを明示する音の出力がなされる。すなわち、本来の合成音に併せて別の音や合成音が出力される。

【0022】

【発明の実施の形態】以下、本発明の実施の形態につき図面を参照して説明する。

【第1の実施形態】図1は本発明の音声合成システムの第1の実施形態を示すブロック構成図である。

【0023】図1のシステムの中心をなす音声合成部11は、入力テキストの言語解析を行う言語解析部111と、その解析結果に対して規則を適用して、音韻的な制御を行う音韻制御部112と韻律的な制御を行う韻律制御部113と、音韻制御部112および韻律制御部113の制御に従い音声波形を生成する波形生成部114と、生成された波形を出力する波形出力部115とから構成されている。この音声合成部11の構成の枠組みに

ついては、既存のテキスト音声変換可能な音声合成システムの一般的な構成法がそのまま利用できる。テキスト音声合成システムの一般的な構成法としては、例えば佐藤他：“日本語テキストからの音声合成，”電気通信研究所研究実用化報告，Vol.32，No.11，pp.2243-2252（1983年11月）などが利用できる。

【0024】音声合成部11内の言語解析部111は、入力テキストに対して形態素解析や統語構造解析などの言語解析を行い、形態素の系列に分解するとともに「読み」を表す記号列と形態素の品詞、活用、アクセント型、形態素間の係り受け関係の強さなどの情報を併せて出力する。

【0025】音声合成部11内の音韻制御部112および韻律制御部113は、これら言語解析部111での言語解析結果の内容に対して、それぞれ様々なレベルの規則を適用することによって、言語メディアから音声メディアへの変換に伴う合成音声の品質を左右する制御を行う。

【0026】具体的には、アクセント単位の認定（すなわち、未知語へのアクセント付与、アクセント結合、複合語のアクセント分割、付属語連鎖に対する副次アクセントの付与などが含まれる）、読みの認定（すなわち、未知語への読み付与、連濁処理、表記から読みへの変換などが含まれる）、1つの韻律的なまとまりとするためのいわゆる韻律語（アクセント句）連鎖のまとまりの認定（すなわち、韻律句境界の付与がこれに相当する）、韻律制御パラメータ値決定（すなわち、ピッチの時間変化パターンを生成するモデルのパラメータ値の決定、音韻・ポーズのタイミングの決定、パワーの決定）、読みに対応する蓄積パターンの検索および蓄積素片の選択（すなわち、蓄積単位への変換、検索条件への変換、複数の検索結果を得たときの選択などが含まれる）、蓄積素片の編集（すなわち、蓄積素片間の接続、補間加工などが含まれる）といった各段階の処理をそれぞれの規則によって行う。

【0027】このうち、音韻制御部112においては、読みの認定、読みに対応する蓄積パターンの検索および蓄積素片の選択、蓄積素片の編集を、韻律制御部113においては、アクセント単位の認定、韻律語（アクセント句）連鎖のまとまりの認定、韻律制御パラメータ値決定を担当する。これらの各段階の処理や規則は任意の分類が可能であり、システムの実装形態によって上記の分類とは異なる場合や省略される場合もあるが、基本的にはこれらの内容に沿った処理が行われる。また、システムの実装形態によっては、前段の言語解析部111や後段の波形生成部114との境界も様々であるが、ここでは、上記の音韻的な制御を行うものとして音韻制御部112を、同じく上記の韻律的な制御を行うものとして韻律制御部113を、それぞれ定義している。

【0028】韻律制御部113は、言語解析部111で

の形態素解析結果にこれらの規則を適用して、形態素系列の読みに対応する個々の音韻やポーズなどのタイミングを決め、形態素系列あるいは読みに対応する音韻の系列を韻律語（アクセント句）というアクセント付与のための韻律制御上の単位に分割するとともに、意味上の文構造上や生理的な制約による呼吸段落上のまとまりを形成し、いわゆる韻律句と呼ばれる話調成分付与のための韻律制御上の単位へ韻律語系列を分割し、各韻律制御上の単位に対して、タイミングを考慮して、アクセントや話調の成分の大きさを与えるパラメータ値を決めピッチを決定する。韻律制御部113はさらに、形態素系列あるいは読みに対応する音韻の系列あるいはピッチなどをもとにパワー包絡を決定する。

【0029】一方、音韻制御部112は、読みに対応する音韻の部分系列に対して、音声波形、あるいは音声波形の分析パラメータ、あるいはその両方を対応させた蓄積素片を格納しておく蓄積データ格納部1121を有しており、この蓄積データ格納部1121に格納されている蓄積素片のバリエーションを考慮して、形態素系列の読みに当たる音韻の部分系列に対応する蓄積素片の系列を決定する。

【0030】本実施形態において、上記した音韻制御部112および韻律制御部113で適用される規則は、計算機内の通信状態や計算機外との間の通信状態に応じて切り替えられるようになっているが、これについては後述する。

【0031】音声合成部11内の波形生成部114は、音韻制御部112の出力する蓄積素片系列を接続し、韻律制御部113の出力する制御情報、すなわち、タイミング、ピッチ、パワー包絡に従い、信号処理レベルでの韻律制御を行って、音声波形を生成する。

【0032】音声合成部11内の波形出力部115は、音声合成部11により生成された音声波形を例えばスピーカーやイヤホン等から出力する。さて、本実施形態において、音韻制御部112が持つ蓄積データ格納部1121に格納される蓄積素片、音韻制御部112で利用される規則、および韻律制御部113で利用される規則は、生成したい合成音声の調子に合う自然音声データを収集しておき、そのデータから予め作成しておいたものである。例えば、対話調の音声を合成したい場合は模擬対話音声を収集したり、ささやき声、早口の声、疲れた様子の音、元気のいい声、雑踏の中で（あるいは雑踏環境を模擬したところで）発声した声、落ち着いた声、様々な人の声をできるだけ大量に収集し、それぞれのピッチやパワー、時間長の分析結果から、それぞれの声に対応した規則や蓄積データを導出する。

【0033】様々な状況における音声が、それぞれ異なる傾向の音韻的・韻律的な特徴を有することは従来からの研究で指摘されており、様々な音声データから導かれた韻律の制御規則が異なる傾向を示すことは、平井他：

“種々の音声コーパスから自動生成されたF₀制御規則の違いについて、” 日本音響学会講演論文集、2-5-3, p.271-272（1994年10月～11月）においても実際のデータとともに示されている。

【0034】音声データからの規則の導出に関しては従来から研究例が多数ある。例えば、広瀬他：“音声合成とアクセント・イントネーション、” 電子情報通信学会誌、Vol.70, No.4, pp.378-385（1987年4月）、三村他：“統計的手法を用いた音声パワーの分析と制御、” 日本音響学会誌、Vol.49, No.2, pp.253-259（1993年12月）、海木他：“発話速度による文音声のポーズ長変化の分析、” 日本音響学会講演論文集、1-5-16, pp.247-248（1992年12月）などがあり、規則の抽出に利用できる。

【0035】それぞれの環境について抽出された制御規則および蓄積素片には、音声合成時に利用するための抽出環境に関する情報、即ち、対話調であるとか、ささやき声、早口の声、疲れた様子の声、元気のいい声、雑踏の中での声、落ち着いた声、などの音声データの収集状況の情報が付加される。

【0036】周知のように、既存の音声合成システムの音韻的・韻律的な制御規則や蓄積データは、本質的には、言語的な環境（例えば、形態素、品詞、活用など）および音韻的・韻律的な環境（例えば、音韻の並び、アクセント型とアクセント核、ピッチ、パワー包絡、タイミングなど）と制御内容（例えば、読み記号列、アクセント結合情報、韻律パラメータ値、蓄積素片の選択優先度など）や音声波形・分析パラメータとの対応として捉えることができる。

【0037】そこで本実施形態では、この対応関係に規則の抽出環境を加え、抽出した制御規則や蓄積データを、言語的な環境、音韻的・韻律的な環境および規則の抽出環境と、制御内容や音声波形・分析パラメータとの対応として記述している。

【0038】このように、複数の規則や蓄積データを備え、さらに、それらを音声合成部11（内の言語解析部111および音韻制御部112）が適宜選択して使用することにより、合成音声の声の調子にバリエーションを与えることができる。

【0039】そこで本実施形態では、上述した音声合成部11に加えて、当該音声合成部11での規則選択の条件を決定するための情報を与える手段として、通信状態監視部12が設けられている。この通信状態監視部12は、音声合成システムが稼働する計算機内の通信状態を監視する計算機内通信状態監視部121と、当該計算機と外部の間の通信状態を監視する計算機外通信状態監視部122とを有している。

【0040】通信状態監視部12内の計算機内通信状態監視部121は、同一計算機内で動作するソフトウェア間、ハードウェア間、あるいはソフトウェアとハードウ

エアの間の通信状況ないしは通信路の品質からなる通信状態を監視する。説明を簡単にするために、ここでは互いに通信を行うハードウェアあるいはソフトウェアをそれぞれ通信者Aおよび通信者Bと簡略化して表現する。すなわち通信者Aと通信者Bとの間で通信が行われているものとする。

【0041】計算機内通信状態監視部121は、これらの間で交わされる通信状態を知るために、この通信を媒介するソフトウェアないしはハードウェア（便宜的にここでは、通信媒体と呼ぶことにする）に問い合わせ、通信状況（例えば、情報の送り手、通信量や通信量の時間的な変化、通信の頻度、送る予定のデータ総量、既にしたデータ量など）や通信路の品質（例えば、データ転送速度やエラー発生頻度など）を通知してもらう。これらの通知は、必ずしも問い合わせが必要なわけではなく、問い合わせがなくても通信媒体側から計算機内通信状態監視部121に適当なタイミングで通知するようにしても構わない。

【0042】このような通信媒体として、オペレーティングシステムやオペレーションシステム（以下、OSと称する）の提供する既存の機能（例えば、メッセージング機能を実現できるWindowsのDDE=Dynamic Data Exchangeや、クリップボードを使ったデータの受け渡し）や、ウィンドウシステムの提供する既存の機能（例えば、X Window Systemにおけるイベントやセレクションバッファ、Windowsのmessageなどが一例である）、あるいは、サーバ・クライアントモデルで実装された各種サービスが利用できる。もちろん、既存システムを利用するだけでなく、同様のメカニズムを持つように新たなシステムを組むことも可能である。

【0043】また、通信者Aと通信者Bで交わされる通信状態を知るために、通信媒体を介さずに直接、通信者Aと通信者Bに問い合わせる仕組みにしてもよい。この場合、通信者Aおよび通信者Bがそれぞれ持っている、通信を行う機能を持つ部分（便宜的に、通信部と呼ぶことにする）に対して計算機内通信状態監視部121が問い合わせ、上記と同様に通信状態を通知してもらう。もちろん、上記と同様に、問い合わせがなくても適宜、通信者Aおよび通信者Bがそれぞれ持っている通信部が計算機内通信状態監視部121に適当なタイミングで通知するようにしても構わない。

【0044】計算機内通信状態監視部121は、このようにして取得した通信状態に関する情報をもとに、例えば、通信量が多い／小さい、送るべきデータ総量が多い／少ない、既に通信が済んだデータの割合が多い／小さい、データ転送速度が速い／遅いといった情報を通信状態情報として音声合成部11に送る。これらの情報は、取得した数値のまま通信状態情報としてもよいし、計算機内通信状態監視部121内で閾値と比較して離散的なレベルにまとめて通信状態情報としてもよい。

【0045】一方、通信状態監視部12内の計算機外通信状態監視部122は、計算機外との通信状態を監視する。この計算機外通信状態監視部122においても、上記した計算機内通信状態監視部121と同様に、通信媒体を介して通信状態を取得する構成とすることができる。通信媒体としては、同じように、OSやOSの提供する既存の機能（メッセージング機能）や、ウィンドウシステムの提供する既存の機能（例えば、イベント）、あるいは、サーバ・クライアントモデルで実装された各種サービス（例えば、Network File Systemやプリンタのデーモン等）の他、モデムのように計算機外とのデータ通信が可能なデバイスやドライバが利用できる。もちろん、既存システムを利用するだけでなく、同様のメカニズムを持つように新たなシステムを組むことも、上記と同様に可能である。

【0046】音声合成部11は（通信状態監視部12内の）計算機内通信状態監視部121および計算機外通信状態監視部122からそれぞれ通信状態情報を受け取り、当該通信状態情報に応じて音韻制御部112および韻律制御部113においてそれぞれ適用する制御規則や蓄積データを選択する。

【0047】ここで、通信状態情報と選択する制御規則および蓄積データとの対応関係は、音韻制御部112および韻律制御部113に定めておく。例えば、通信量大あるいは通信の頻度が大きい場合は早口にしたり、非常に大きい場合には緊迫した声にしたり、逆に、通信量少あるいは通信の頻度が小さい場合は、ピッチのダイナミックレンジを大きく、落ち着いた声にしたり、ポーズを多めに挿入したり、ゆったりした声にする。通信残量が多い場合はのんびりした声に、残り少なくなってくるにつれ、ピッチを高めにしたり早口にしたりする。通信路の品質が悪い場合には、元気のない声やピッチに不規則な揺らぎを重畳させ声質を変える。転送速度が速ければ軽快な声を、遅ければ重苦しい声にするなどの対応関係が例として挙げられる。

【0048】このように、言語解析部111での解析結果に対して音声合成部11（内の言語解析部111および音韻制御部112）において適用する制御規則や蓄積データを、通信状態監視部12（内の計算機内通信状態監視部121または計算機外通信状態監視部122）から出力される通信状態情報（の示す通信状態）に応じて切り替えて（変更して）合成音声出力することにより、ユーザは、合成音声の声の調子から、その時点における計算機内の通信状態、あるいは計算機外との間の通信状態を知ることができる。

【0049】なお、上述の対応関係はあくまで例であって、音声合成システムのユーザの好みに応じて変更可能にしても構わない。また、計算機内通信状態監視部121で監視される計算機内の通信状態と、計算機外通信状態監視部122で監視される計算機外の通信状態のそれ

それぞれについて、独立に対応関係を設定しても構わない。

【0050】また、以上の実施形態では、通信状態監視部12には、計算機内通信状態監視部121および計算機外通信状態監視部122の両方が設けられているものとしたが、いずれか一方だけが設けられているものであっても構わない。

【第2の実施形態】図2は本発明の音声合成システムの第2の実施形態を示すブロック構成図である。なお、図1と同一部分には同一符号を付してある。

【0051】まず、図2の構成の特徴は、音声合成部1に加えて、計算機ハードウェアの動作状態を監視するハードウェア状態監視部221と計算機ソフトウェアの動作状態を監視するソフトウェア状態監視部222とを有する動作状態監視部22が設けられている点である。これに伴い、図2における音声合成部11内（の音韻制御部112および韻律制御部113）の機能も、以下に述べるように図1中の音声合成部11（内の音韻制御部112および韻律制御部113）とは異なるが、便宜上同一符号を付してある。

【0052】動作状態監視部22内のハードウェア状態監視部221は、音声合成システムの稼働する計算機ハードウェアの動作状態を示すパラメータを直接測定したり、あるいは、計算機ハードウェアもしくはそのソフトウェアドライバに動作状態を問い合わせたり、あるいは、計算機ハードウェアもしくはそのソフトウェアドライバ自体から適当なタイミングで動作状態を通知されることによって、計算機ハードウェアの動作状態を監視する。

【0053】例えば、システムを構成するハードウェアに供給される電源電圧の高さや安定性、カード、プリンタ、キーボード、マウス等のデバイス（周辺機器）やネットワークケーブル等、システムに接続されているハードウェアの接続状況（接続されているか否か、さらには利用可能な状態か否か）を監視する。

【0054】ハードウェア状態監視部221は、このようにして取得したハードウェア状態に関する監視結果をもとに、例えば、電源電圧が十分高い／高い／やや低い／低い／かなり低い、十分安定している／安定している／やや不安定／非常に不安定、などにランク分けされる電源品質に関する情報、あるいは、ハードウェアが利用可能な状態にある／待機状態にある／接続が切れているといった動作状態情報を音声合成部11に送る。

【0055】なお、上記の分類は一例であり、必要に応じて任意の分類が可能である。また、適当な閾値を設定し、これと比較して離散的なレベルにまとめてもよいし、取得した数値のまま動作状態情報としてもよく、上記の分類に限定されるものではない。

【0056】音声合成部11は（動作状態監視部22内の）ハードウェア状態監視部221から動作状態情報を受け取り、当該動作状態情報に応じて音韻制御部112

および韻律制御部113においてそれぞれ適用する制御規則や蓄積データを選択する。

【0057】ここで、動作状態情報と選択する制御規則および蓄積データとの対応関係は、前記第1の実施形態における通信状態情報と選択する制御規則および蓄積データとの対応関係と同様に、音韻制御部112および韻律制御部113に定めておく。この対応関係は、例えば、品質の高い電源電圧が十分安定して供給されている場合は通常の韻律制御や声色で合成音声を生じさせるが、電源電圧が下がり始めたり不安定な場合には、少し元気がない声に対応する蓄積データを選択するような規則を選択したり、ゆったりした口調になるような規則に切り替えたり、ピッチの上げ下げを弱めたりするような規則を選択したり、おとなしい声になるような規則を選択したりするなどの対応関係が例として挙げられる。もちろん、この対応関係はあくまで例であって、音声合成システムのユーザの好みに応じて変更可能にしても構わない。そして、これらの規則の対応関係の変更は、上記と逆の印象を与えるように選択であっても構わない。

【0058】音声合成部11内の韻律制御部113および音韻制御部112では、このような対応関係に従って選択された規則を用いることで、生成・出力する合成音声の韻律的・音韻的な品質を制御する。これによりユーザは、合成音声の声の調子から、その時点における計算機ハードウェアの状態を知ることができる。

【0059】ところで、PDA（Personal Digital Assistants）に代表される可搬型システム（携帯機器）では、表示に利用できる面積が小さいことから、システムの動作状態情報を提示するために割り当てる面積を大きくするのは非効率的であるが、過度に小さくすればユーザに注意を促すという本来の目的が満足できなくなる可能性がある。一般に、可搬型のシステムの場合、供給される電源の安定性は、整備された環境にある固定型のシステムに比較して低いのが普通である。そこで、可搬型システムにおいて、このような動作状態情報について、韻律や声質を制御することによって副次的に伝えることは有効である。電源電圧と同様、ハードウェアの接続状況も、一般のユーザが普段は比較的意識せず、見落としがちであるが、これも接続状況の変化に応じて韻律や声質に変化を与えれば、ユーザにそれとなく知らせることができる。

【0060】一方、動作状態監視部22内のソフトウェア状態監視部222は、音声合成システムの稼働する計算機のプロセッサ（CPU）やメモリ、ハードディスク等の計算機資源をある（ターゲットとする）ソフトウェア（プロセス）がどれだけ占有しているか、あるいは、逆の観点から言えば、あるソフトウェアがどれだけ処理を待たされているかといった、限られた計算機資源の分配に起因するソフトウェアの動作状態を監視したり、あるソフトウェアが今どのような入力を受け付ける状態に

あるか、例えば、入力デバイスの種類や入力内容の種類として何が有効であるか、また、あるソフトウェアがどのような情報を提示しているか、例えば、提示情報の出所や提示内容の種類といった動作モード（場面）に対応するソフトウェアの動作状態を取得する。

【0061】このようなソフトウェアの動作状態は、ソフトウェア状態監視部222が、当該ソフトウェアの動作しているOSに対して問い合わせ、通知してもらうことによって取得したり、当該ソフトウェア自体に、動作状態を直接問い合わせると通知する通知部（通知機能）を付加しておくことによって取得する。もちろん、問い合わせがなくても、ソフトウェア自身がその動作状態をソフトウェア状態監視部222に適当なタイミングで通知する仕組みを用意することも可能である。

【0062】ここで、取得するソフトウェアの動作状態情報としては、例えばメモリ使用量やソフトウェア状態、CPUの占有率や占有時間累計、動作優先度等の情報が一例として挙げられる。これらの情報は、既存のOSのシステムコールやライブラリを利用して取得可能である。また、現在受け付ける入力の種類や提示している情報の種類を通知する通知部を備えたソフトウェアを新たに作成してもよい。

【0063】一般に、同一のアプリケーションでも動作モード（場面）に応じて受け付ける入力の種類は動的に変化する。例えば、メールの送受信を行うメールアプリケーションは、届いたメールのリストを表示する状態、そのうちの選択されている1つのメールの内容を表示する状態、送信したいメールの文面を編集する場面、編集したメールを送信する場面などがあって、それぞれの場面によって、同じキー入力が有効になるか無視されるか、有効であった場合にどのような動作をするかが変わってくる。また、音声認識入力を受け付けるソフトウェアの場合には、今どのような認識語彙が入力可能であるかといった情報が「受け付けられる入力の種類」に相当し、さらに、認識語彙だけでなくそれぞれの認識語彙に対応する動作もソフトウェアの動作モード（場面）に応じて動的に変化する。

【0064】一方、電子メールのアプリケーションでは、誰から送られたメールであるとか、極秘扱いの内容であるといった、情報の出所や内容を表す情報を文字列照合や言語解析によって取得し、これらの動作モードや提示情報の出所や提示内容の種類を動作情報としてソフトウェア動作状態監視部222に伝える。ここでは、メールのアプリケーションを例に挙げたが、電子ネット掲示板や電子ネット上の情報提供システムのように、複数の情報源からの情報をブラウジングする応用ソフトウェアにおいても全く同様のことが適用できる。

【0065】ソフトウェア状態監視部222は、このようにして取得したソフトウェアの動作状態に関する情報をもとに、例えば、メモリの占有が大きい／小さい、C

PUの占有時間累計が長い／短い、認識語彙の組合せがどのセットであるか、どのような動作モードにあるか、情報の出所はどこか、情報の内容の種類が何であるかを示す情報を動作状態情報として音声合成部11（内の音韻制御部112および韻律制御部113）に送る。

【0066】音声合成部11では、ソフトウェア状態監視部222からの動作状態情報を受け取ると、音韻制御部112および韻律制御部113においてそれぞれ適用する規則や蓄積データを当該動作状態情報に応じて選択する。これにより、例えばメモリの占有が大きいとか、CPUの占有時間累計が長い場合には、元気のいい声や申し訳なさそうな声を生成して、システムの状態をユーザにそれとなく伝えたり、逆に早口の口調とすることでユーザ自らの処理を促したりすることが可能となる。また、情報の出所に応じて、アクセントやフレーズを変えるための規則を選択し、地域色を音声に反映することで、情報の出所の違いをユーザに意識させることが可能となる。また、情報提供者の声の蓄積データがあれば、それを使うことで、情報提供者を簡単に判別できるようにすることも可能である。また、電話等でリモート操作する場合や、携帯機器で表示面積が小さい場合に、あるソフトウェアが現在どのような入力を受け付ける状態にあるか（その入力デバイスの種類と入力内容の種類）に応じて、韻律や声色に変化を与えることで、ユーザは次に何を入力すべきかや、現在の「場面」を、出力される合成音声の調子から知ることができる。

【0067】ここで、動作状態情報（ソフトウェアの動作状態情報）と選択する制御規則および蓄積データの対応関係は、上述した計算機ハードウェアの動作状態情報と選択する制御規則および蓄積データの対応関係の場合と同様に、音韻制御部112および韻律制御部113に定めておく。

【0068】このように、言語解析部111での解析結果に対して音声合成部11（内の音韻制御部112および韻律制御部113）において適用する制御規則や蓄積データを、動作状態監視部22（内のハードウェア状態監視部221またはソフトウェア状態監視部222）から出力される動作状態情報に応じて切り替えて（変更して）合成音声を出力することにより、ユーザは、合成音声の調子から、その時点における計算機ハードウェアの状態、あるいは計算機ソフトウェアの状態を知ることができる。

【0069】さて、本実施形態における音声合成部11では、言語解析部111、音韻制御部112、韻律制御部113、波形生成部114、および波形出力部115のそれぞれが単独で動作するようにモジュール化しておき、互いのデータの授受の形式がネットワークを通じたものであっても、同一の実行プロセス内でのデータの授受でも処理が可能になるようにしてある。また、上記各部の全体の処理手続きおよびその部分的な処理手続きが

互いに別のプロセスとして分離可能にしておき、分離されたプロセスは、処理結果を元のプロセスに返すようにしておく。このようなシステムの実装は、マルチタスクOS上ならば、子プロセスの生成と子プロセスとのソケット通信などのシステムコール、ライブラリを用いて容易に実装可能である。

【0070】音声合成部11は、動作状態監視部32から動作状態情報を受け取ると、メモリの残量やCPUの占有時間や占有率から判断して、当該音声合成部11を構成する言語解析部111から波形出力部115に至る音声合成処理を進めるに当たってメモリやCPU能力など十分な計算機資源が確保されているか否かをチェックする。そして音声合成部11では、メモリが不足する可能性がある場合や、CPUの負荷状況から十分な計算機資源が確保できないと判断される場合には、現在までに処理が進んでいる段階よりも後の処理のうち適当なものを別の計算機ハードウェアに別プロセスとして分担させ、処理結果を受け取るようにする。

【0071】ここで、どの処理を分担させるかは、処理に必要なCPU能力やメモリ量から判断するが、これは音声合成方式の種類や蓄積データの規模によって変わるものである。例えば、分析パラメータ合成方式の場合には、波形生成部114における信号処理、次いで音韻制御部112における蓄積データの編集加工の処理に資源が多く使われ、波形素片編集型の音声合成方式では、蓄積データの検索がCPUパワーを最も要する。分析パラメータ合成方式においても、蓄積データが持つ蓄積素片の種類が多いほど蓄積データの検索時間は多くかかる。したがって、どの処理を分担させるかは、合成方式や蓄積データの規模によって適当な優先順位をつけて、それに従って分担させる処理を決めればよい。

【0072】ところで、音声合成部11の言語解析や音韻・韻律的な制御における規則の適用は、いずれも多くの規則適用の可能性を数え上げ、その規則を適用した場合の結果を評価することによって、処理が進められる。これらの処理を1つの計算機上で順次実行するのは必ずしも効率的でなく、規則の適用可能性と規則適用を仮定した場合の評価を同時並行して処理する方が効率的である。適用される規則が固定的であれば、ある程度チューニングすることにより順次処理をさせることによって問題は比較的顕れにくくすることも可能であるが、本発明のように適用される規則が動的に変更される場合には、同時並行的な処理をする方が効率的である。

【0073】そこで本実施形態では、上述の規則の適用可能性と規則適用を仮定した場合の評価を同時並行して処理するようにしている。この並行処理は、計算機が接続されたネットワーク上の他の計算機にリモートプロセスとして実行させたり、同じ計算機上の副プロセッサに分担させることもできる。

【0074】このように本実施形態においては、音声合

成部11を構成する言語解析部111から波形出力部115に至る音声合成処理を進める上で、動作状態監視部32からの動作状態情報により十分な計算機資源が確保できないと判断される場合には、現在までに処理が進んでいる段階よりも後の処理のうち適当なものを別の計算機ハードウェアに別プロセスとして分担させたり、言語解析や音韻・韻律的な制御における規則の適用可能性と規則適用を仮定した場合の評価を、計算機が接続されたネットワーク上の他の計算機や同じ計算機上の副プロセッサに分担させて同時並行して処理させることで、効率的な処理を実現し、ユーザの待ち時間を減らすようにしている。

【0075】なお、以上の実施形態では、動作状態監視部22には、ハードウェア状態監視部221およびソフトウェア状態監視部222の両方が設けられているものとしたが、いずれか一方だけが設けられているものであっても構わない。

【第3の実施形態】図3は本発明の音声合成システムの第3の実施形態を示すブロック構成図である。なお、図1と同一部分には同一符号を付してある。

【0076】まず、図3の構成の特徴は、音声合成部11に加えて、ユーザのシステム利用状況を監視する利用者状態監視部321と、ユーザのシステム利用環境を監視する利用者環境監視部322とを有するユーザ状況監視部32が設けられている点である。これに伴い、図3における音声合成部11内（の音韻制御部112および韻律制御部113）の機能も、以下に述べるように図1中の音声合成部11（内の音韻制御部112および韻律制御部113）とは異なるが、便宜上同一符号を付してある。

【0077】ユーザ状況監視部32内の利用者状態監視部321は、ユーザのシステムの利用状況（利用者状態）を得るための入力デバイスや時計、利用履歴の少なくとも1つからの情報を監視し、例えばユーザがどの程度集中してシステムを利用しているかといったシステム利用状況監視結果を取得する。入力デバイスとしては、例えば、カメラなどが利用できる。カメラの捉えたユーザの頭の向きを精度よく推定することは可能であり、ある一定時間中にどの程度長く安定してシステムの方向（正面）を向いているのか否か（他を向いているか）をもって、ユーザの集中度として評価する。また、マウスに代表されるポインティングデバイス、キーボード等、ユーザの入力操作のための入力デバイスについて、ユーザの操作状況（入力操作頻度、入力操作時間、ポインティングデバイス移動速度・距離など）を監視することも可能である。この他、時計、利用履歴については、同じ曜日、同じ時間帯にどのような利用状況にあったかを記録しておくことで、利用状況の推定精度を高めるのに用いられる。

【0078】利用者状態監視部321は、このようにし

て取得したユーザのシステム利用状況に関する情報をもとに、ユーザの集中度や、ユーザの操作状況を示す情報をユーザ状況情報として音声合成部11（内の音韻制御部112および韻律制御部113）に送る。

【0079】音声合成部11では、利用者状態監視部321からのユーザ状況情報を受け取ると、音韻制御部112および韻律制御部113においてそれぞれ適用する規則や蓄積データを当該ユーザ状況情報に応じて選択する。これにより、例えば集中度が予め定めた閾値以下の場合には、パワーを大きくしたり、文頭では発話速度が小さく（遅く）なるような規則を適用することで、ユーザに集中するように注意を促すことが可能となる。

【0080】一方、ユーザ状況監視部32内の利用者環境監視部322は、ユーザがシステムを利用している場所の環境（利用者環境）を得るための入力デバイスや時計、利用履歴の少なくとも1つからの情報を監視し、例えばユーザがどのような音環境（周囲雑音環境）下に居るかとか、どの程度の明るさの場所に居るかとか、ユーザの物理的な居場所（位置）といったシステム利用環境監視結果を出力する。このような入力デバイスとして、例えば、周囲雑音を集音するマイクロフォンや、GPSなどの位置推定デバイス、さらには明るさセンサ、カメラ、ガスセンサ、水センサなどが挙げられる。その他、時計、利用履歴については、同じ曜日、同じ時間帯にどのような利用環境にあったかを記録しておくことで、利用環境の推定精度を高めるのに用いられる。

【0081】利用者環境監視部322は、このようにして取得したユーザのシステム利用環境に関する情報をもとに、周囲雑音のスペクトル特徴やレベル、明るさ、ユーザの居場所（位置）等を示す情報をユーザ状況情報として音声合成部11（内の音韻制御部112および韻律制御部113）に送る。

【0082】音声合成部11では、利用者環境監視部322からのユーザ状況情報を受け取ると、音韻制御部112および韻律制御部113においてそれぞれ適用する規則や蓄積データを当該ユーザ状況情報に応じて選択する。これにより、例えば高周波数成分に優勢な雑音がある場合には、はっきり聞こえるように高いピッチの声になるように韻律制御規則を適用したり高周波数成分の優勢な蓄積素片を選択するように音韻制御規則を適用するとか、雑音レベルが低い静かなところでは、静かな声あるいは落ち着いた声になるような規則を適用することができる。また、明るい場所で利用する際はピッチが高めで発話速度が早くなるように、暗い場所で利用する際には発話速度を遅く、ピッチのダイナミックレンジは広くなるように韻律規則を適用することで、明るい場所に比較して暗い場所では比較的落ちついた印象を与えることもできる。このような対応関係は、ユーザの好みに応じて変更可能としても構わない。

【0083】なお、以上の実施形態では、ユーザ状況監

視部32には、利用者状態監視部321および利用者環境監視部322の両方が設けられているものとしたが、いずれか一方だけが設けられているものであっても構わない。

【第4の実施形態】図4は本発明の音声合成システムの第4の実施形態を示すブロック構成図である。なお、図3と同一部分には同一符号を付してある。

【0084】まず、図4の構成の特徴は、図3の構成（の音声合成部11およびユーザ状況監視部32）に加えて、非自然音声明示判断部41と、テキスト変更部42が設けられている点である。これに伴い、図4における音声合成部11（内の言語解析部111等）の機能も、以下に述べるように図3中の音声合成部11（内の言語解析部111等）とは異なるが、便宜上同一符号を付してある。

【0085】まず非自然音声明示判断部41は、ユーザ状況監視部32の出力するユーザ状況情報をもとに、人間の発声でないこと（非自然音声であること）を明示すべきか否かを判断し、その判断結果（非自然音声明示判断結果）を出力する。例えば、ユーザ状況情報においてユーザが集中していないことを示している場合や、これまであまり合成音声の出力をしたことのない時間帯や場所であることを示している場合には、人間の発声でないことを明示すべきであるという判断結果を出力する。

【0086】テキスト変更部42は非自然音声明示判断部41からの非自然音声明示判断結果を受け取り、当該判断結果が人間の発声でないことを明示すべきことを示している場合には、入力テキストに対応する合成音声の出力に先立ち（すなわち、言語解析部111での入力テキストに対する言語解析結果を音韻制御部112および韻律制御部113に出力して、対応する音声波形を生成・出力させるのに先立ち）、音声合成によるメッセージ出力が始まることを予告する「合成音です」「システムからのお知らせです」などの定型表現を前置する。音声合成部11は、このテキスト変更部42によって前置された語彙を含めて合成出力する。

【0087】このようにして、例えばユーザが集中していない場合や、これまであまり合成音声の出力をしたことのない時間帯や場所での利用の場合に、音声合成によるメッセージ出力が始まることを予告する（非自然音声であることを明示する）合成音を、入力テキストに対応する合成音声の出力に先立って出力することで、そのような状況をユーザに知らせることができる。特に、高品質で肉声に近い合成音声出力される状況では、人の声がする利用者環境のもとでの利用の場合に、非自然音声であることを明示する合成音を前置することで、周囲の人の声と紛らわしくしないとか、非自然音声であることを明示しないことで、合成音であることを強調して注意を集めるのを避けることが可能である。

【第5の実施形態】図5は本発明の音声合成システムの

第5の実施形態を示すブロック構成図である。なお、図4と同一部分には同一符号を付してある。

【0088】まず、図5の構成の特徴は、図4で示したテキスト変更部42に代えて、人間の発声でないことを明示する音（非自然音声明示音）を出力する非自然音声明示音出力部43が設けられている点である。これに伴い、図4における音声合成部11の機能も、例えば当該音声合成部11内の波形出力部115が、波形生成部114により生成される合成音声と、非自然音声明示音出力部43の生成する非自然音声明示音とを混合する機能を有しているというように、図4中の音声合成部11（内の波形出力部115等）とは異なるが、便宜上同一符号を付してある。

【0089】まず、非自然音声明示音出力部43は、非自然音声明示判断部41から出力される非自然音声明示判断結果が人間の発声でないことを明示すべきことを示している場合には、入力テキストに対応する合成音声の出力に先立ち、例えば「ピ」といった信号音（非自然音声明示音）を出力する。この信号音は、音韻制御部112および韻律制御部113による音韻・韻律的な制御に従って波形生成部114により生成される合成音声の出力に先立ち、波形出力部115によって出力される。

【0090】このようにして、ユーザが集中していない場合や、これまであまり合成音声の出力をしたことのない時間帯や場所での利用の場合に、例えば「ピ」という非自然音声明示音を、入力テキストに対応する合成音声の出力に先立って出力することで、人間の発した声ではなく合成音声によるメッセージであることを明示してユーザに対して注意を促すことができる。

【0091】なお、図5の構成に図4中のテキスト変更部42を加え、このテキスト変更部42と非自然音声明示音出力部43の両方を備えた構成とすることも可能である。

【第6の実施形態】図6は本発明の音声合成システムの第6の実施形態を示すブロック構成図である。なお、図1と同一部分には同一符号を付してある。

【0092】まず、図6の構成の特徴は、図1の構成（の音声合成部11および通信状態監視部12）に加えて、図4に示したような入力テキストの変更を行うテキスト変更部42が設けられている点である（但し、テキスト変更の内容が、図4の例とは異なる）。これに伴い、音声合成部11内（の言語解析部111等）の機能も、以下に述べるように図1中の音声合成部11（内の言語解析部111等）とは異なるが、便宜上同一符号を付してある。

【0093】図6の構成の音声合成システムにおいて、音声合成部11内の言語解析部111は、通信状態監視部12から通信状態情報を受け取ると、当該情報をテキスト変更部42に渡して起動する。

【0094】するとテキスト変更部42は、言語解析部

111と連絡をとりながら、言語解析部111により言語解析されている入力テキストに通信状態情報に応じた定型表現の語彙を挿入して当該テキストを変更する。すなわちテキスト変更部42は、音声合成部11内の音韻制御部112および韻律制御部113の処理の先頭において、あるいは、韻律制御部113の処理の途中においてポーズ挿入位置を決めた段階において、文頭や文末、あるいはポーズ挿入位置に、通信状態情報によって決まる定型表現の語彙を挿入する。音声合成部11は、このテキスト変更部42によって挿入された語彙を含めて合成出力する。

【0095】以上のテキスト変更部42での通信状態情報に応じたテキスト変更処理により、例えば、通信量大のとき（通信が混んでいるとき）には、「あ」「えーと」「えー」「はい」などの不要語を文頭や文末、あるいはポーズ挿入位置に挿入したり、「ちょっと待って」などのメッセージを文頭に前置することができる。このような決まった語彙（あらかじめ設定されている語彙）を挿入することによって、処理時間をかせぎ合成音声の処理による負荷を低減する効果がある。また、逆に通信量小のときには、上記と同様の不要語を挿入すれば、システムがアイドル状態であることをユーザにそれとなく示すという効果がある。

【0096】なお、図6の構成におけるテキスト変更部42は、通信状態監視部12からの通信状態情報を音声合成部11を通して受け取るものとしているが、通信状態監視部12から直接受け取るようにしても構わない。

【第7の実施形態】図7は本発明の音声合成システムの第7の実施形態を示すブロック構成図である。なお、図2と同一部分には同一符号を付してある。

【0097】まず、図7の構成の特徴は、図2の構成（の音声合成部11および動作状態監視部22）に加えて、図6に示したようなテキスト変更部42が設けられている点である。これに伴い、図7における音声合成部11内（の言語解析部111等）の機能も、以下に述べるように図2中の音声合成部11（内の言語解析部111等）とは異なるが、便宜上同一符号を付してある。

【0098】図7の構成の音声合成システムにおいて、音声合成部11内の言語解析部111は、動作状態監視部22からシステムの動作状態情報を受け取ると、当該情報をテキスト変更部42に渡して起動する。

【0099】するとテキスト変更部42は、言語解析部111と連絡をとりながら、言語解析部111により言語解析されている入力テキストに動作状態情報に応じた定型表現の語彙を挿入する。すなわちテキスト変更部42は、音声合成部11内の音韻制御部112および韻律制御部113の処理の先頭において、あるいは、韻律制御部113の処理の途中においてポーズ挿入位置を決めた段階において、文頭や文末、あるいはポーズ挿入位置に、動作状態情報によって決まる定型表現の語彙を挿入

する。音声合成部11は、このテキスト変更部42によって挿入された語彙を含めて合成出力する。

【0100】以上のテキスト変更部42での動作状態情報に応じたテキスト変更処理により、例えば、CPUが長時間占有されているときには、「あ」「えーと」「えー」「はい」などの不要語を文頭や文末、あるいはポーズ挿入位置に挿入することができる。このような決まった語彙を挿入することによって、処理時間をかせぎ音声の処理による負荷を低減する効果がある。

【0101】なお、図7の構成におけるテキスト変更部42は、動作状態監視部22からの動作状態情報を音声合成部11を通して受け取るものとしているが、動作状態監視部22から直接受け取るようにしても構わない。

【第8の実施形態】図8は本発明の音声合成システムの第8の実施形態を示すブロック構成図である。なお、図3と同一部分には同一符号を付してある。

【0102】まず、図8の構成の特徴は、図3の構成（の音声合成部11およびユーザ状況監視部32）に加えて、図6に示したようなテキスト変更部42が設けられている点である。これに伴い、図8における音声合成部11内（の言語解析部111等）の機能も、以下に述べるように図3中の音声合成部11（内の言語解析部111等）とは異なるが、便宜上同一符号を付してある。

【0103】図8の構成の音声合成システムにおいて、音声合成部11内の言語解析部111は、ユーザ状況監視部32からユーザ状況情報を受け取ると、当該情報をテキスト変更部42に渡して起動する。

【0104】するとテキスト変更部42は、言語解析部111と連絡をとりながら、言語解析部111により言語解析されている入力テキストにユーザ状況情報に応じた定型表現の語彙を挿入する。すなわちテキスト変更部42は、音声合成部11内の音韻制御部112および韻律制御部113の処理の先頭において、あるいは、韻律制御部113の処理の途中においてポーズ挿入位置を決めた段階において、文頭や文末、あるいはポーズ挿入位置に、ユーザ状況情報によって決まる定型表現の語彙を挿入する。音声合成部11は、このテキスト変更部42によって挿入された語彙を含めて合成出力する。

【0105】以上のテキスト変更部42での動作状態情報に応じたテキスト変更処理により、例えば、ユーザが集中していないときには、「あの」などの人に声をかける語彙を文頭に設定することで、ユーザに注意を促すことができる。

【0106】なお、図8の構成におけるテキスト変更部42は、ユーザ状況監視部32からのユーザ状況情報を音声合成部11を通して受け取るものとしているが、ユーザ状況監視部32から直接受け取るようにしても構わない。

【0107】

【発明の効果】以上詳述したように本発明によれば、言

語情報の持つメッセージとしての直接的な情報伝達だけでなく、音声合成機能を含む、システム全体の状況をそれとなく示す、音声メディアの持つ副次的な情報伝達機能を利用し、使い勝手のよいシステム構築が可能になる。また、ユーザの利用状況に応じた合成音声出力が可能となる。

【0108】特に、計算機の出力量メディアとして、システム内部の状態をユーザに伝えることはユーザインタフェースの観点からも重要である。言語メッセージ伝達としての主目的としての利用と同時に、システムがどのような動作状況にあるかをそれとなくユーザに伝えることは、音声メディアの利用形態として適切なものであるといえる。

【0109】このような情報は画面表示部などの視覚的な出力と併用することでその効果を高めることも可能であるだけでなく、PDAに代表される携帯機器のように表示部の面積が小さい場合には、メッセージを、主に音声メディアによって伝えるようにすれば、メッセージ表示による画面の面積の占有を抑えることが可能になる。

【0110】さらに、ユーザの利用状況を考慮して韻律や声色を制御することにより、より自然なシステム出力が可能になる。それは、状況を考慮せず単調な合成音声を出力しないようにするというだけでなく、高品質で肉声に近い合成音声がもらえるような状況では、逆に合成音声であることを明らかにして、音声としては不自然さはあっても、機械とのコミュニケーションとしては自然なやりとりが可能になる。

【図面の簡単な説明】

【図1】本発明の音声合成システムの第1の実施形態を示すブロック構成図。

【図2】本発明の音声合成システムの第2の実施形態を示すブロック構成図。

【図3】本発明の音声合成システムの第3の実施形態を示すブロック構成図。

【図4】本発明の音声合成システムの第4の実施形態を示すブロック構成図。

【図5】本発明の音声合成システムの第5の実施形態を示すブロック構成図。

【図6】本発明の音声合成システムの第6の実施形態を示すブロック構成図。

【図7】本発明の音声合成システムの第7の実施形態を示すブロック構成図。

【図8】本発明の音声合成システムの第8の実施形態を示すブロック構成図。

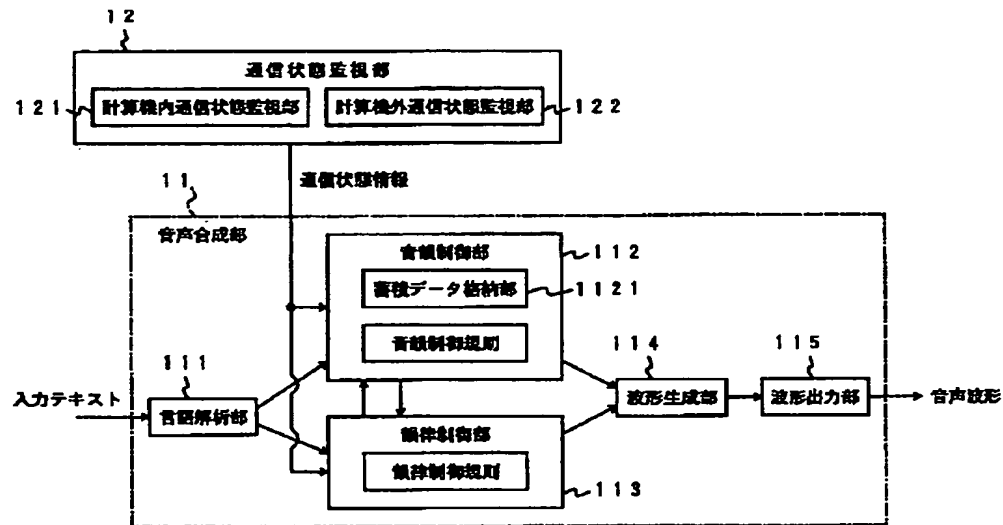
【符号の説明】

- 11…音声合成部、
- 12…通信状態監視部、
- 22…動作状態監視部、
- 32…ユーザ状況監視部、
- 41…非自然音声明示判断部、

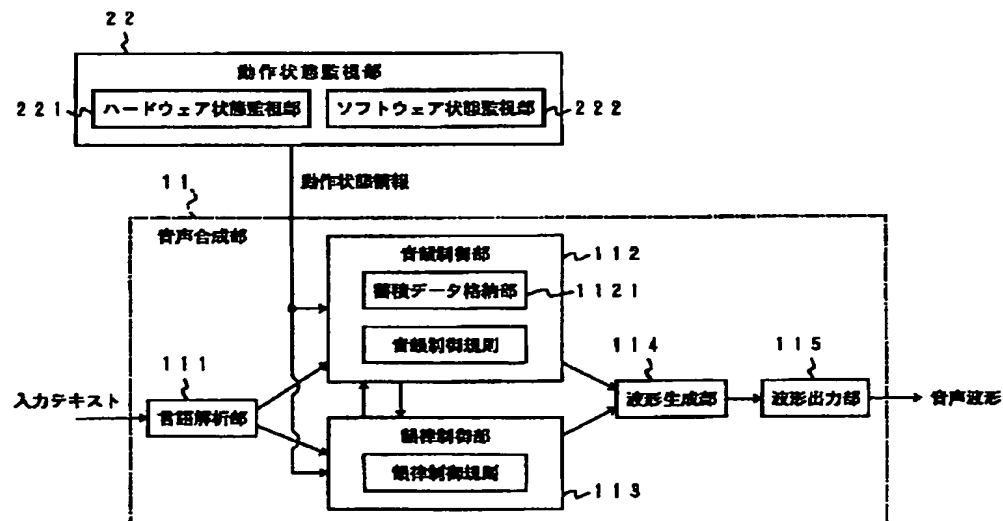
4 2 …テキスト変更部、
 4 3 …非自然音声明示音出力部、
 1 1 1 …言語解析部、
 1 1 2 …音韻制御部、
 1 1 3 …韻律制御部、
 1 1 4 …波形生成部、
 1 1 5 …波形出力部、

1 2 1 …計算機内通信状態監視部、
 1 2 2 …計算機外通信状態監視部、
 2 2 1 …ハードウェア状態監視部、
 2 2 2 …ソフトウェア状態監視部、
 3 2 1 …利用者状態監視部、
 3 2 2 …利用者環境監視部、
 1 1 2 1 …蓄積データ格納部。

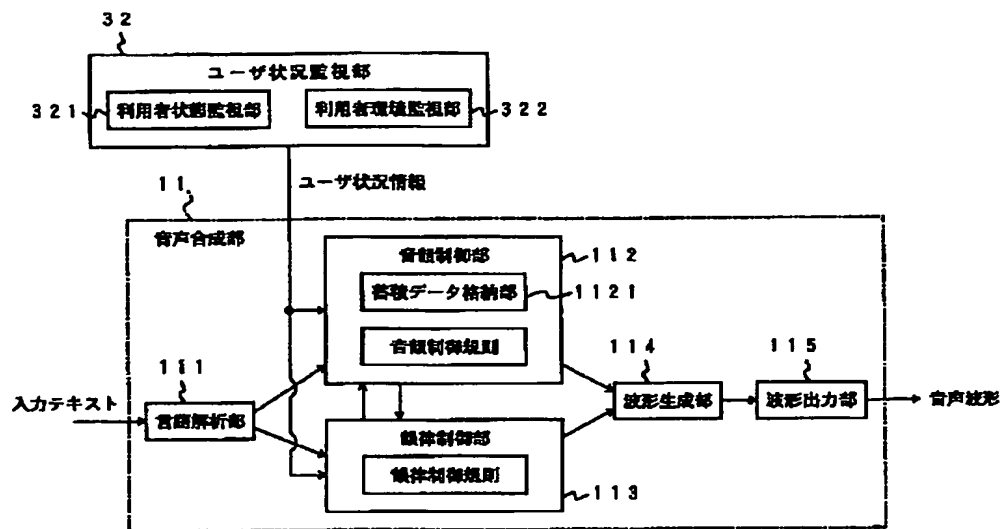
【図 1】



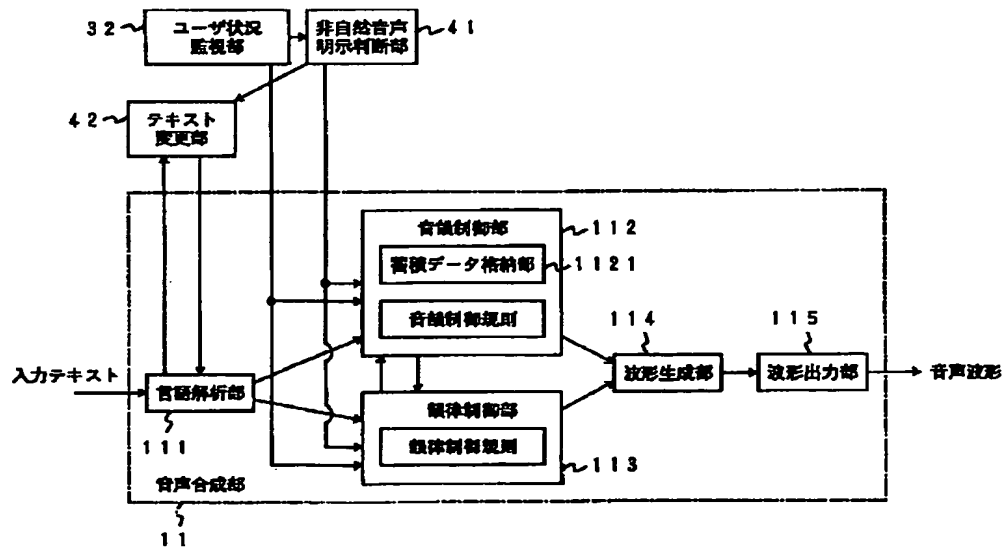
【図 2】



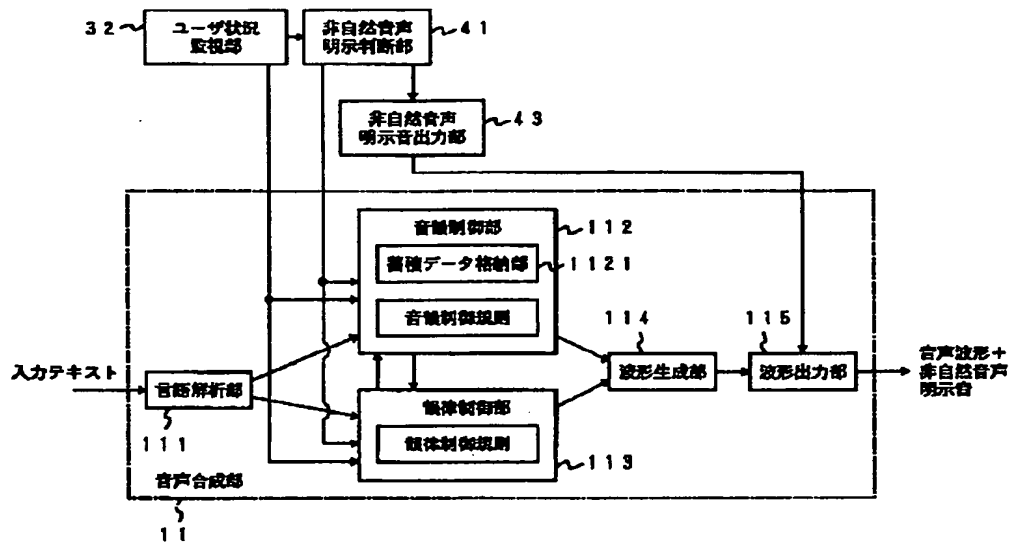
【図3】



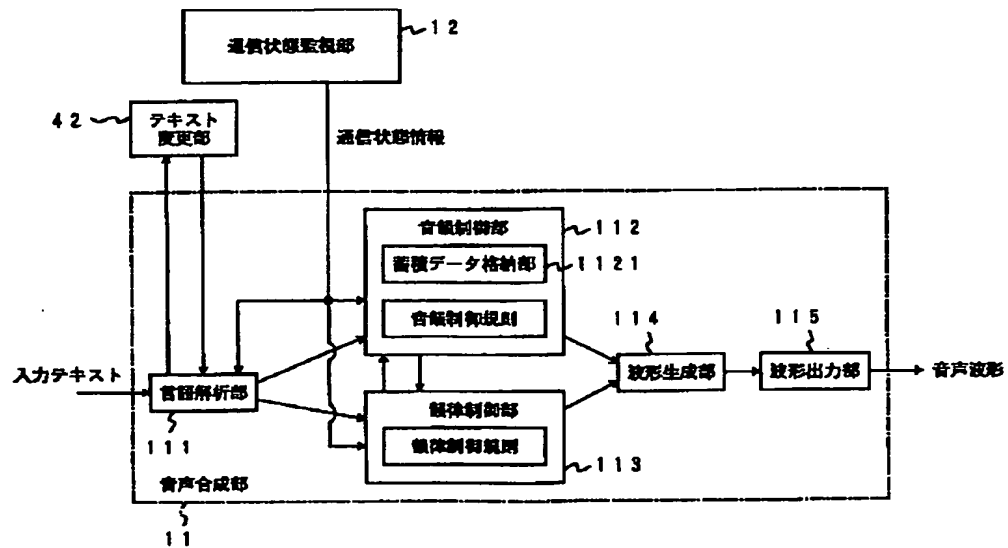
【図4】



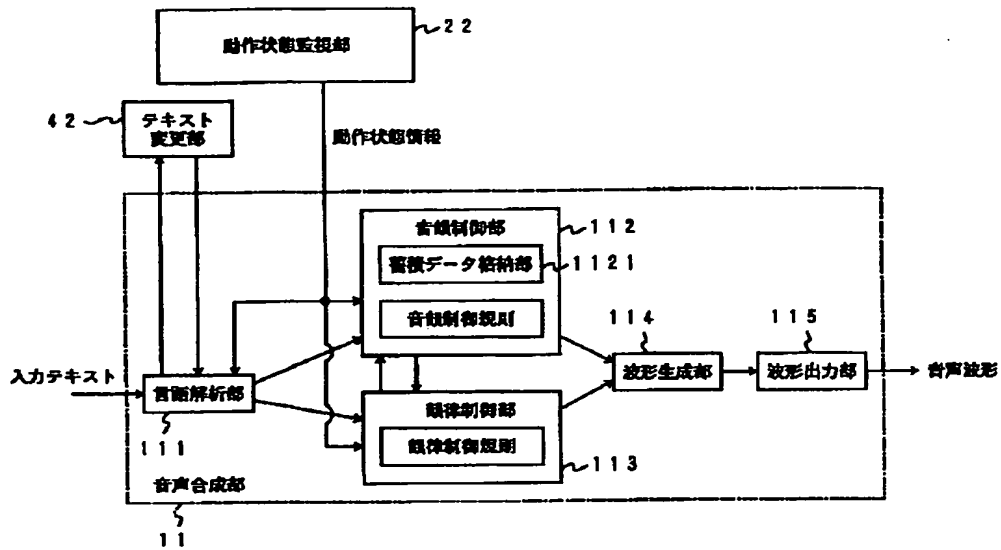
【図5】



【図6】



【図7】



【図8】

